

# **РАЗРАБОТКА СИСТЕМЫ СТАТИСТИЧЕСКОГО АНАЛИЗА СООБЩЕНИЙ В СОЦИАЛЬНЫХ СЕТЯХ С ПРИМЕНЕНИЕМ МОДЕЛИ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ MAPREDUCE**

**Игнатенко С. А.**

*ИООО "ЭПАМ Системз", Минск, Беларусь,  
e-mail: ihnatsenka@gmail.com*

Широкое распространение параллельных архитектур вычислительных систем вызывает повышенный интерес к средствам разработки программного обеспечения, способного максимально полно использовать аппаратные ресурсы данного типа. Однако к текущему моменту имеется определенный разрыв между имеющимися на потребительском рынке технологиями аппаратной реализации параллелизма и программными средствами их поддержки. Так, если многоядерные компьютеры общего назначения стали нормой в середине предыдущего десятилетия, то появление OpenMP, популярного стандарта разработки программ для подобных систем, отмечено еще почти десятью годами ранее. Практически в то же время возник и стандарт MPI [1], описывающий способы передачи сообщений между процессами в распределенной среде. Развитие обоих данных стандартов, выражающееся только в расширении функциональности без адаптации парадигм к объектно-ориентированному подходу, приводит к тому, что они оказываются несовместимы с современными платформами программирования, такими как Microsoft .NET Framework [2]. Поэтому разработчикам этих платформ приходится прилагать дополнительные усилия по внедрению средств параллелизма в свои продукты.

На основе модели распределенных вычислений MapReduce в контексте платформы Nadoop с последующей реализацией на Microsoft .NET Framework создана уникальная динамическая библиотека для статистического анализа текста и разработано трехуровневое приложение (клиент, сервер, облако), которое позволяет выполнить статистический анализ текста сообщений в социальных сетях.

Отличительной особенностью методики облачных вычислений с использованием алгоритма MapReduce на основе Microsoft .NET фреймворка является высокая производительность и скорость работы, что приводит к экономии системных ресурсов и уменьшению материальных расходов при выполнении аналогичной задачи с помощью других методов. Практической значимостью работы является возможность использования ее результатов для успешного решения задач, требующих анализа большого объема текста. Области применения полученных результатов могут являться: выявление сложных зависимостей в социальных сетях, выявление мошеннических действий в финансовом секторе, анализ логов и другие.

Практические результаты по созданию платформы представлены в виде презентации.

## **Литература**

1. The Message Passing Interface (MPI) standard [Electronic resource]. – Mode of access: <http://www.mcs.anl.gov/research/projects/mpi>.